

From Radio Design to System Evaluations for Ultra-Reliable and Low-Latency Communication

Shehzad Ali Ashraf*, Y.-P. Eric Wang*, Sameh Eldessoki[†],
Bernd Holfeld[†], Donald Parruca[‡], Martin Serror[§], James Gross[¶]

*Ericsson Research: {shehzad.ali.ashraf, eric.y.p.wang}@ericsson.com

[†]Fraunhofer Heinrich Hertz Institute: {sameh.eldessoki, bernd.holfeld}@hhi.fraunhofer.de

[‡]Paderborn University: donald.parruca@upb.de

[§]COMSYS, RWTH Aachen University: martin.serror@comsys.rwth-aachen.de

[¶]School of Electrical Engineering, KTH Royal Institute of Technology: james.gross@ee.kth.se

Abstract—Ultra-reliable and low-latency communication is the enabler for many new use cases, including wireless industrial automation. Fulfilling varying requirements of these use cases demands a flexible radio design. To address this, a holistic approach needs to be adopted. Therefore, this paper presents the radio access concepts affecting the communication reliability and latency, and comprehensively evaluates link and system level considerations through simulations. In particular, we describe the choice of suitable modulation and coding schemes, and discuss the impact of different numerologies and waveform candidates. We also point out the key principles for radio frame design to reduce the end-to-end latency. The presented concepts are then used to evaluate the performance at system level for an industrial scenario. It is shown that by an appropriate design of the radio interface for 5G system, the required low-latency and high reliability for industrial applications and many other use cases can be achieved.

Keywords—5G; URLLC; Industrial IoT; Wireless industrial automation; M2M; Radio interface design; Low latency; High reliability; Industrie 4.0.

I. INTRODUCTION

Wireless connectivity has seen major advancements and evolutions in recent years. Under the label of Internet-of-Things (IoT), new use cases with varying requirements have emerged and are categorized as either Massive Machine-Type Communication (mMTC) or Ultra-Reliable and Low-Latency Communication (URLLC) [1], targeting different services and markets. While mMTC targets massive deployments of low cost devices with extended battery lifetime, URLLC aims to support applications requiring extremely low Packet Loss Rate (PLR) and highly reduced end-to-end latency. Among the URLLC use cases are some new market segments such as industrial real-time controlled automation, intelligent transport systems and tactile internet. However, this paper particularly focusses on wireless industrial automation which is the most challenging use case in terms of latency and reliability requirements.

Currently, 3GPP defines the 5G URLLC requirements as a reliability of $1 - 10^{-5}$ for the transmission of a small packet with a user-plane latency of 1 ms [2]. However, when specifically looking in the domain of industrial automation systems, the constraints on PLR can be as low as 10^{-9} [3], assuming short packet sizes of a few bytes. Also depending on the application, the latency requirement may vary from larger

than 10 ms for data acquisition and monitoring applications down to 0.5 ms-1 ms for motion control and alarms [4]. Therefore, to address the vertical market of industrial automation, a flexible communication system, targeting the most challenging requirements of reliability (i.e. 10^{-9} PLR) and latency (i.e. < 1 ms), needs to be designed.

While a number of recent research articles [5], [6], [7] discuss requirements and practical implementation challenges of URLLC, a holistic perspective towards the radio design is missing. One reason is that the transmission reliability and latency need to be tackled with methods on various layers of the communication protocol stack.

In this paper, we introduce different radio design concepts and integrate them into one comprehensive framework that approaches both link and system level. In principle, the paper extends our previous work in [8], [9] by considering further aspects to achieve reliable communication with strict latency constraints. As the industrial channel characteristics play a pivotal role in radio link optimization, the presented schemes also take account of the propagation measurements [10], [11] at sub-6 GHz spectrum.

Our main contributions in the next sections are as follows:

- 1) Discussion on methods to achieve reliable communication. (cf. Section II)
- 2) Selection of Forward Error Correction (FEC) and physical layer abstraction to be used for system level evaluations. (cf. Section III)
- 3) Discussion on the impact of waveform numerologies and signaling channels placement on achievable communication latency. (cf. Section IV)
- 4) Performance of different waveforms for mixed numerologies. (cf. Section V)
- 5) System level performance for industrial network layout. (cf. Section VI)

II. METHODS TO ACHIEVE RELIABLE COMMUNICATION

Primary drivers of the radio design for URLLC system are: (i) the latency requirement, which is translated into the transmission time interval, (ii) the reliability requirement, which is interpreted as the targeted residual Block Error Rate (BLER) requirement, and (iii) the size of data packets. The packet size and the transmission time together determine the required

data rate, which in turn requires a certain Signal-to-Noise Ratio (SNR) for a successful decoding. In fading channel environments, an additional power margin is needed to fulfill the BLER requirement. The power margin depends on the probability that the whole channel fades down and can be reduced by using different diversity techniques, both on lower and higher layers of the protocol stack. In [9], authors suggest the use of diversity order up to 16 to get the feasible radio link budget for highly reliable communication. From a radio design perspective, there are several ways to exploit such high diversity orders such as time diversity, frequency diversity and spatial diversity. However, time diversity is not considered as an attractive option for URLLC due to restrictions on latency. In this section, we focus on exploiting spatial diversity to combat the fading characteristics of the propagation channel. Please note that frequency diversity can be exploited on top of spatial diversity given sufficient spectrum availability.

One of the most popular spatial diversity techniques is to use multiple antennas at the wireless transmitter and/or receiver. Transmit diversity can be achieved with space-frequency block coding such as applying Alamouti coding over multiple transmit antennas across a frequency resource pair. On the receiver side, multiple antennas not only achieve a diversity gain, but also provide a receiver processing gain from coherently combining the received signals from multiple receive antennas. However, exploiting spatial diversity by solely relying on multiple antennas might not always be possible. For instance, when the transmission devices have only a single antenna such as small sensor devices in industrial scenarios. Also, the spatial correlation between different antennas on such small devices affects the achievable diversity gains.

Hence, to further increase the reliability of the wireless transmission system for URLLC, additional techniques exploiting macro-diversity can also be considered on the higher layers of communication protocols. Multi-connectivity concepts as specified in 3GPP can be used to achieve macro-diversity. That means, duplicating the packet transmission over different communication links to enhance the probability of successful packet reception. Furthermore, other techniques to exploit macro-diversity can also be used but their benefit lies in an appropriate network structure. As shown in [3], an appropriate architecture design can also allow for network-assisted Device-to-Device (D2D) communication. That means, the central controller (or the base station) locally assigns transmission resources to the connected devices, which then directly communicate to each other avoiding an additional hop via the base station. Nevertheless, such transmissions are locally broadcasted within the radio cell due to the nature of radio communication. This redundancy can also be leveraged to further increase the reliability of the system. More specifically, locally overhearing devices can be used as a relay to retransmit packets over a presumably uncorrelated transmission path. This special case of spatial diversity is often referred as *cooperative diversity* [12], where devices cooperatively form a virtual antenna array. Recent research showed the use of cooperative diversity specifically for URLLC

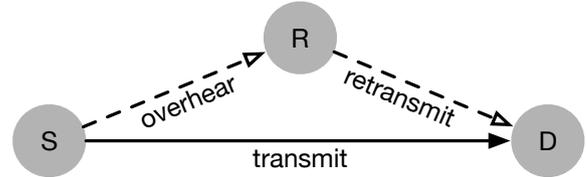


Figure 1: Basic cooperative ARQ scenario: Source (S) transmits to destination (D) a packet, which is overheard by relay (R). If the packet does not reach D, R retransmits it to D.

in industrial control scenarios [13], [14] to achieve the required high reliability. However, it does not consider the timing and signaling overheads which are also crucial design aspects of URLLC. In the following, we take into account the timing and overhead constraints, and propose a new scheme that efficiently exploits cooperative diversity.

A well-known technique to leverage cooperative diversity is cooperative Automatic Repeat reQuest (ARQ), where only unsuccessful transmissions are retransmitted by a relay. Figure 1 depicts the basic cooperative ARQ (cARQ) scenario, including a source (S), a destination (D), and a relay (R). When a packet transmission from S to D fails, R, which overhears local transmissions, retransmits the packet to D thus leveraging a presumably uncorrelated transmission path. Therefore, we also propose the integration of cARQ using network-assisted D2D communication, as a complimentary scheme on top of the physical layer reliability mechanisms considered in the remainder of this paper. In particular, we allow the base station to overhear transmissions between devices and eventually retransmit those packets that did not reach their destination. A benefit of centrally managing retransmissions is the reduced signaling overhead, which is favorable for URLLC [15]. However, when assuming only average Channel State Information (CSI) at the base station, it is not known a priori which transmissions successfully reach their destination and which do not. This implies that a sophisticated mechanism for ACKs / NACKs needs to be implemented to not waste transmission resources and consequently forfeit transmission reliability. Acquiring instantaneous CSI at the base station allows pre-scheduling of transmissions via relay paths according to the currently best link qualities [16]. With each additional relay option, e.g., multiple antennas at the base station or other devices in the network, the diversity order increases [12] and consequently the target reliability levels can be achieved. Nevertheless, this implies that instantaneous CSI about all relevant links needs to be centrally collected at the base station, which might again affect the available transmission resources.

Based on the system model of [16], Figure 2 shows the analytic performance results for the outage probability ϵ for instantaneous CSI at the base station, depending on the CSI signaling overhead T_{csi} . In this model, the signaling overhead does not influence the quality of the instantaneous CSI, which is always assumed to be perfect. Therefore, a case where the base station has only average CSI is additionally considered as a reference. For instantaneous CSI, the outage probability

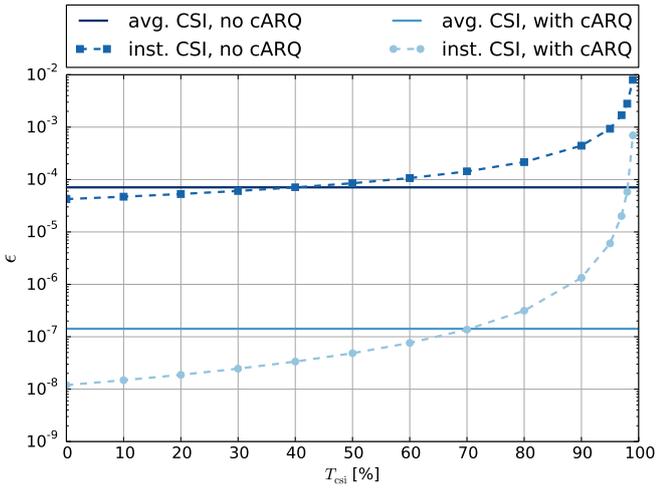


Figure 2: Outage probability ϵ depending on the CSI signaling overhead T_{csi} (considered system model based on [16]).

increases with the signaling overhead as available transmission resources decrease. However, when implementing cARQ and limiting the signaling overhead to $T_{\text{csi}} \leq 20\%$, the reliability gains an additional order of magnitude compared to cARQ with average CSI. Thus, when integrating cARQ into the radio design, it should be considered that with instantaneous CSI at the base station the outage probability potentially goes down, while the signaling overhead might increase, which can again significantly reduce the performance. Therefore an optimum working point is to be determined to limit the CSI overhead and at the same time leverage the cARQ gains. Furthermore, to achieve the anticipated reliability goals, additional diversity techniques as discussed earlier can also be considered.

III. ERROR CORRECTION CODES FOR URLLC

FEC codes are typically used in communication systems to control errors in data transmission over fading channels. In addition, coding is also significant to harvest the gains from diversity techniques mentioned in the previous section. There exist in literature many families of FEC codes offering different coding gains and implementation complexities. Currently, Turbo codes are used in 3GPP systems for data channels and can already meet the physical layer BLER target down to 10^{-5} . However, Convolutional Code (CC) is considered in [9] to reduce the BLER down to 10^{-9} . It is mainly due to their inherent advantages of lacking the error floor as compared to the more advanced FEC codes such as the Turbo codes. Furthermore, convolutional codes perform similarly for very short packet sizes of URLLC applications. A special form of CC is called Tail-Biting Convolutional Code (TBCC) which does not need extra bits to bring the encoder state to the all-zero state. Still Tail-Terminating Convolutional Code (TTCC) were considered in [9] based on the concerns that TBCC offers higher decoding complexity and latency. In the following, we first compare the performance of TBCC as traditionally used for 3GPP LTE with two different families of TTCC and

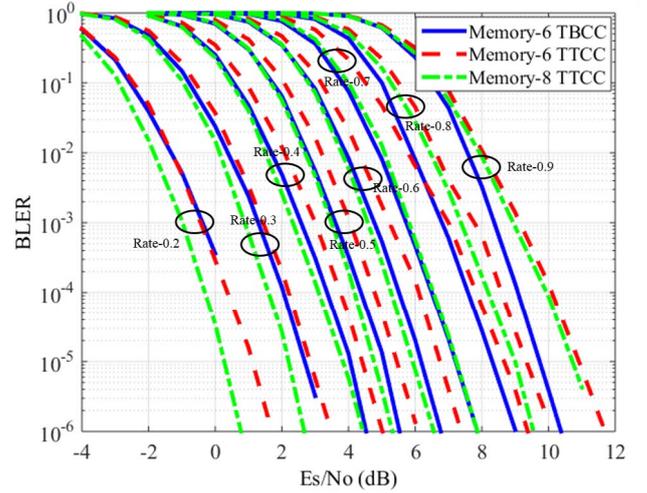


Figure 3: Performance of the considered code families using 100 bits with QPSK modulation in AWGN channel.

show that it performs best. Secondly, we show that the low-complexity decoder can be used for TBCC without impairing its performance and hence, eliminating the doubt of high decoding delay.

1) Link-level performance of tail-biting convolutional codes:

We compare herein the performance of three CC families:

- *Memory-6 TBCC* based on a rate-1/3 mother code with the generator polynomial given as [117 121 91].
- *Memory-6 TTCC* based on rate-1/4, 1/3 and 1/2 with the generator polynomials given as [121 117 91 79], [109 87 79] and [109 79], respectively.
- *Memory-8 TTCC* based on rate-1/4, 1/3 and 1/2 with the generator polynomials given as [441 421 315 351], [381 403 471] and [285 431], respectively.

All the other code rates in a code family are obtained by either repetition or puncturing. The first two families of convolutional codes are selected to compare the performance of TTCC with TBCC using the same memory length. Furthermore, higher memory TTCC is also compared to analyse its coding gains over the memory-6 TBCC.

Figure 3 shows the link level performance in AWGN channels assuming a data packet size of 100 bits and QPSK modulation. E_s represents the energy per QPSK symbol and N_o is the noise power spectral density. We can observe that memory-8 TTCC family has the best performance for the code rates 0.6 and lower, whereas memory-6 TBCC outperforms for the code rates 0.7 and higher. The reason that memory-6 TBCC outperforms memory-8 TTCC is because it is easier to optimize the puncturing pattern for obtaining a high-rate TBCC. Furthermore, memory-6 TTCC family has the worst performance among the three code families for all code rates.

To get further insights into the Modulation and Coding Scheme (MCS) for URLLC and form the basis for link-to-system modeling, we compare different MCSs in fading channels for the discussed families of convolutional codes. We consider four modulation schemes i.e. QPSK, 16-QAM, 64-QAM and 256-QAM, each of which is combined with code

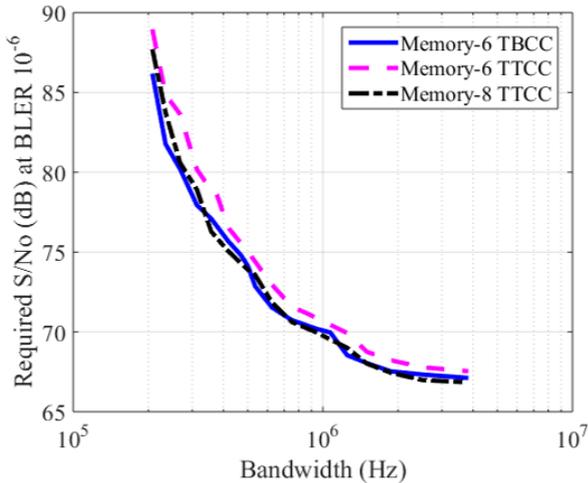


Figure 4: Performance of the considered code families of convolutional codes using 100 bits with antenna diversity in a fading channel.

rate 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9 to give a total of 32 MCSs. As suggested in [9], we use eight transmit antennas and two receive antennas to exploit spatial diversity gains at the physical layer. Please note that the code symbols are spread across all transmit antenna pairs. With this approach, if the minimum free distance is greater than the total diversity order, the code word error probability will benefit from all the diversity channels [17]. Besides, maximum ratio combining is assumed for combining signals from multiple receive antennas without considering correlation between antennas. Figure 4 shows the required SNR-per-Hertz along with the required bandwidth for the best MCSs for the three considered FEC families. It can be observed in Figure 4 that the TBCC has better or similar performance as compared to the other two code families considered, making it a suitable candidate for URLLC. Please note that the simulation results shown in the figures give the required SNR only down to 10^{-6} BLER as the numbers of realizations used in the simulations are in the order of millions. To achieve the SNR at BLER of 10^{-9} , the presented results can be extended using extrapolation methods since convolutional codes do not have an error floor.

2) Low-latency decoder for tail-biting convolutional codes:

As mentioned previously, the decoding complexity of TBCC can be higher than that of TTCC which affects the achievable end-to-end latency. Here, we analyse the complexity of the TBCC decoder and compare the performance of low-latency decoding with the more complex decoding used in 3GPP LTE system as a baseline.

Wrap-around Viterbi decoding techniques on the circular tail-biting trellis are generally employed to decode TBCC. In [18], [19], it is shown that 2x wrap-around is sufficient to achieve good decoding performance. However, 2x wrap-around means that the decoding operation is extended by one full trellis at the end of the original trellis. Given the very tight latency constraints for URLLC, this solution can be problematic especially when the packet size is of a few hundred bits. In the

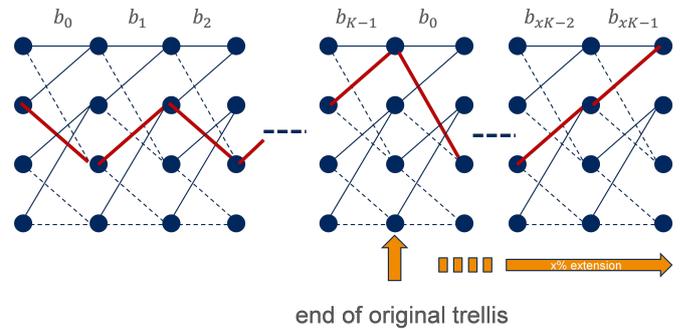


Figure 5: Extend the decoding trellis by $x\%$ of the full trellis.

following, we consider a low-latency TBCC decoder described in [20], [21]. In essence, this low-latency decoder works with any level of trellis extension. At the end of the extended trellis, the trace-back path starts from the ending state that has the best accumulated decoding metric. The ending state does not have to be the same as the starting state on the circular trellis. The decoded bits are taken from the middle of the trace-back path and a circular shift is applied to restore the right ordering of the decoded bits. Figure 5 illustrates the trellis extension in $x\%$ and the trace-back path (represented in red), which starts from the state that has the best decoding metric at the end of the extended decoder trellis. The decoded bits along the trace-back path are identified, where x is the trellis extension factor and K is the length of the information block. $xK/2$ decoded bits on both ends of the extended trellis are discarded, giving K decoded bits. Afterwards, a proper circular shift gives the decoded bits in the right order.

The performance of memory-6 TBCC based on small trellis extension to reduce the decoding delays is shown in Figure 6 considering the packet size of 1000 bits. Please note that 1000 bits packet size is considered here in contrast to 100 bits to see the worst case performance loss for low-latency TBCC decoder. In Figure 6a, performance for a rate-0.3 code is shown and it can be observed that the performance achieved by 5% trellis extension is very close to that achieved by using the baseline decoder with 100% extension. Besides, no sign of error floor is observed. However, the level of trellis extension required to guarantee similar performance as that of a baseline case is also dependent on the code rate. Our evaluations show that 8% trellis extension is needed for a rate-0.6 code as any lower extension factor may result in noticeable performance degradation and error flooring (see Figure 6b). For code-rate 0.9, we see that the extension factor may need to be as high as 25% as shown in Figure 6c. Note also that the level of trellis extension needed is expressed as a percentage of full trellis which in turn determines the number of extra trellis decoding stages needed affecting the overall decoding complexity and latency. The results in Figure 6 suggest that the numbers of extra stages needed are 50, 80 and 250 for rate 0.3, 0.6 and 0.9 codes, respectively. Hence, the decoding complexity and the latency of TBCC can be significantly decreased by the use of low-latency decoders for URLLC without compromising on

the coding gains.

Based on the discussion above, we propose the radio interface design on TBCC for industrial applications requiring a BLER level of 10^{-9} .

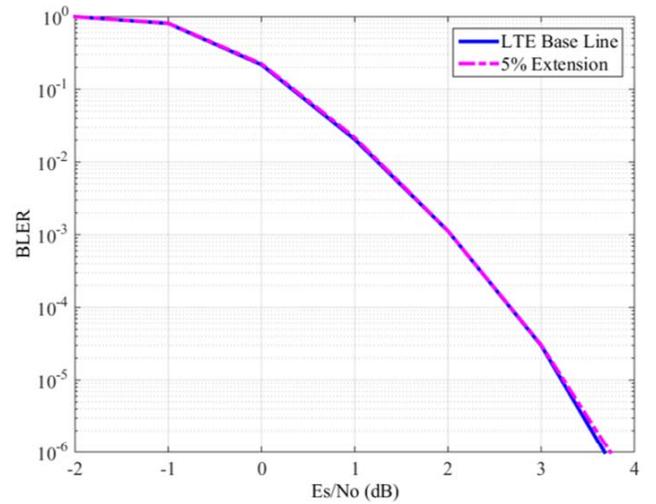
IV. NUMEROLOGIES AND FRAME STRUCTURE DESIGN FOR URLLC

As mentioned in Section I, enabling low-latency communication is similarly important as ensuring transmission reliability for industrial applications. Low-latency communication can be enabled by reducing the data transmission time, and allowing faster signaling and channel access. Therefore, waveform numerology (i.e. transmission symbol size) and frame structure are two critical radio design components in this respect. Below, we discuss the relevant aspects of both related to industrial URLLC systems.

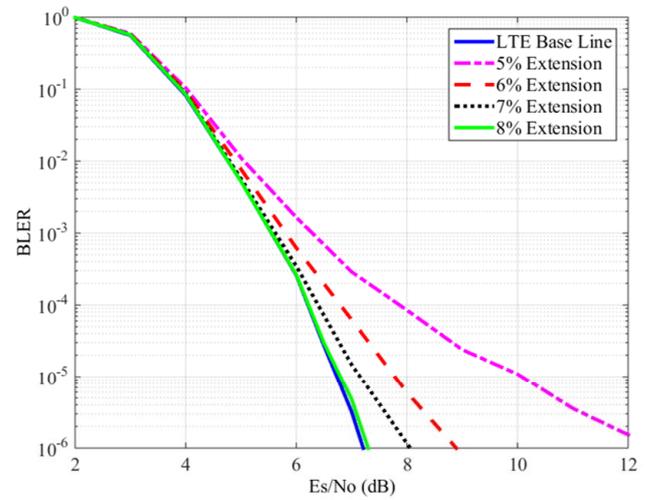
Various factors such as the required Cyclic Prefix (CP) length and the subcarrier spacing Δ_{sc} affect the selection of a particular numerology for a certain application in a particular deployment. The length of CP is dependent on the delay spread of the channel to reduce the effect of inter-symbol interference. On the other side, higher CP duration increases the overhead per transmission symbol. Therefore, it is important to find a good compromise between the CP length and the offered overhead. To allow low-latency communication, shorter transmission symbol size is required which can be achieved using higher subcarrier spacing. At the same time, the required CP overhead puts the upper limit on the choice of subcarrier spacing. Also subcarrier spacing needs to be smaller than the channel coherence bandwidth to allow accurate channel estimation. In short, the choice of subcarrier spacing should be selected as small as possible but still robust against phase noise and Doppler, and at the same time fulfilling the latency requirements. Therefore, a single waveform numerology cannot fulfill the performance requirements of different industrial applications with different deployment scenarios. Hence, a family of numerologies is proposed as shown in Table I, targeting different latency requirements and deployment options.

In Table I, we consider 15 kHz as a baseline numerology and all other numerologies are related to the baseline numerology by an integer scaling factor. For instance, scaling the baseline numerology with a factor of four gives the subcarrier spacing of 60 kHz and symbol size of $16.67 \mu\text{s}$ (excluding the CP). Our measurement campaign in typical indoor industrial environment shows that the delay spread is still much lower than the CP duration of the 60 kHz numerology [10]. Furthermore, we have also observed that there is not much difference in the measured delay spread for different carrier frequencies (e.g. 2.4 GHz and 5.85 GHz).

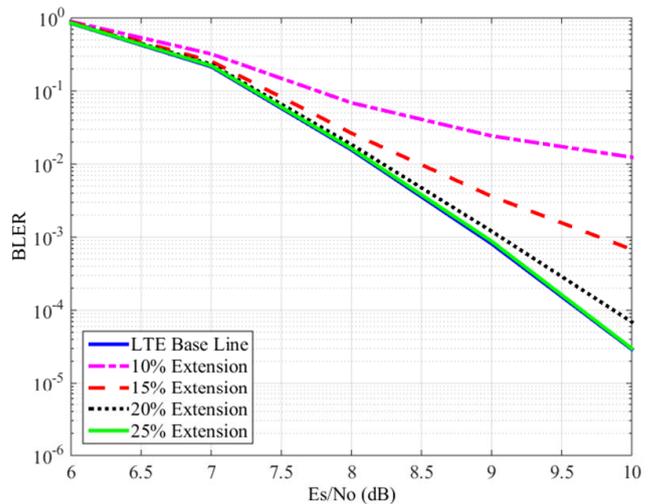
In addition, the radio frame structure should be designed in a flexible way considering different requirements related to industrial applications. Length of data transmission (i.e., transmission time interval), multiplexing of data and control signals, and placement of demodulation reference signals are some of the factors which influence the overall achievable



(a) Code rate 0.3



(b) Code rate 0.6



(c) Code rate 0.9

Figure 6: Tail-biting convolutional code performance with $x\%$ trellis extension and packet size of 1000 bits.

Subcarrier spacing (Δ_{SC})	15 kHz	30 kHz	60 kHz
Symbol duration (T_s)	66.77 μ s	33.33 μ s	16.67 μ s
CP duration (T_{CP})	4.69 μ s	2.35 μ s	1.17 μ s
Clock frequency (f_s)	30.72 MHz	61.44 MHz	122.8 MHz
Symbol samples (N)	2048	2048	2048

Table I: Waveform numerologies relevant for URLLC [22]

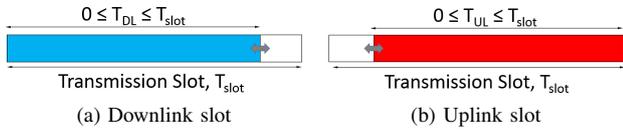


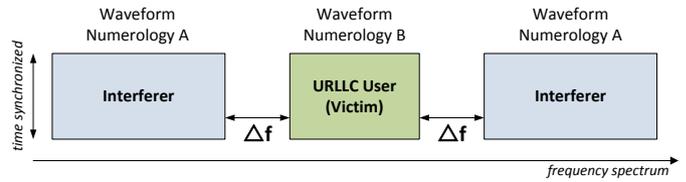
Figure 7: Frame structure enabling low-latency communication

latency of the system. For low-latency communication, it is important to enable fast demodulation and decoding of the received data packet, which requires that the receiver can actually start the processing as early as possible. To enable this, reference (or pilot) signals should be transmitted before the actual data transmission and time-domain interleaving of data should be avoided. With such early decoding, it should be possible for a receiver to demodulate and decode the transmission just a few microseconds after the transmission ended. Moreover, the data and the control signals should be multiplexed in a manner that it allows fast control signaling which reduces the queueing delays.

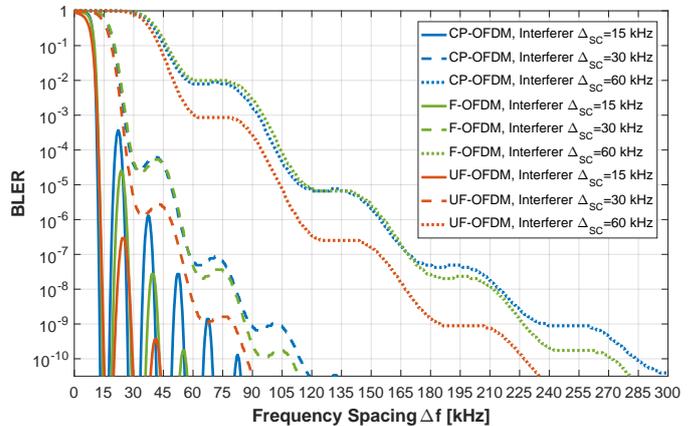
Based on the described design principles, we propose two types of transmission slots as illustrated in Figure 7. T_{slot} is the slot size consisting of a certain number of symbols. The data transmission may not necessarily span all the symbols in the slot by allowing the transmission of Downlink (DL) slot T_{DL} to terminate before the last symbol of the slot and Uplink (UL) transmission T_{UL} to start after the first symbol of the slot. Such slot configuration allows the possibility of having fast UL and DL control signaling within the same slot, respectively. Faster control signaling, whether it is for scheduling the data transmission or transmitting ACKs / NACKs by the receiver, results in lower end-to-end latency. Also note that the same design principles are valid for sidelinks if network-assisted D2D mode is to be used.

V. WAVEFORM IMPACT ON THE COEXISTENCE OF MIXED NUMEROLOGIES FOR URLLC

Varying requirements and deployment scenarios for different industrial applications may result in the use of different numerologies. Therefore, industrial URLLC devices using the same waveform numerologies will not only coexist among each other but also with the devices operating with different numerologies in same carrier. This demands for an efficient use of scarce spectrum resources. In this section, we analyse the trade-off point with respect to the frequency spacing Δf , where out-of-band emission of adjacent interferers becomes too strong to achieve the required communication reliability for the desired URLLC user. Fig. 8a illustrates the considered system model, where the two adjacent interferers have a symmetric frequency spacing. The coexistence trade-off point is affected by several system parameters, e.g., the selection



(a) Considered time-frequency localization for a URLLC user transmitting in the guard band of two adjacent interferers using different waveform numerology (assumption of symmetric frequency spacing Δf).



(b) BLER depending on spacing Δf for the scheme in Fig. 8a when all users transmit on a resource block size of 12 subcarriers and 64-QAM transmission with rate-1/2 convolutional coding is considered (packet size: 100 bits). Numerology modes are mixed between the desired user ($\Delta_{SC} = 15$ kHz) and the interferers ($\Delta_{SC} = \{15, 30, 60\}$ kHz).

Figure 8: Coexistence scheme with mixed OFDM-based waveform numerologies.

of an appropriate waveform, the mix of numerologies, and the occupied transmission bandwidth of both interferers and desired URLLC users.

Due to good properties in time localization and flexibility to enable latency-optimized symbol shapes, waveforms based on Frequency Division Multiplexing (FDM) are in favor for reliable low-latency communication. Orthogonal Frequency Division Multiplexing (OFDM) is by far the most prominent waveform for modern wireless systems but there are various OFDM variants being considered as waveform candidates for 5G systems [23]. In the following, we evaluate the coexistence impact of the common Cyclic Prefix OFDM (CP-OFDM) and its variants such as Filtered OFDM (F-OFDM) and Universal Filtered OFDM (UF-OFDM). The frequently discussed Filter Bank Multi-Carrier (FBMC)/Offset Quadrature Amplitude Modulation (OQAM) waveform family, which considers multi-carrier modulation with a well designed prototype filter, is not in our focus. Although it omits the CP and allows very good frequency localization through strong power decay, its large signal spread in time and implementation complexity make it currently unsuitable for URLLC applications with strict latency constraints.

Due to its construction by Discrete Fourier Transform (DFT), CP-OFDM offers strong time localization with slow power decay in frequency domain, which follows the progress of a

sinc-function. In practice, F-OFDM is used to overcome the slow power decay, enabling lower interference to neighboring frequency bands within the required spectrum masks. Therefore, F-OFDM applies a filter over the complete band. For link level evaluations, we consider the length of the time domain filter (here: Hanning) such that it incurs the same overhead caused by the CP in CP-OFDM. In contrast, UF-OFDM divides the complete band into multiple sub-bands, each having an equal number of subcarriers. These sub-bands are filtered independently, increasing the power decay in frequency domain at the cost of softer time localization in comparison to F-OFDM. Typically, the main properties of CP-OFDM are kept at the expense of an increased computational complexity. Also for UF-OFDM, the length of the time domain filter (here: Dolph-Chebyshev) is chosen such that it incurs the same overhead caused by the CP in CP-OFDM.

Fig. 8b gives simulation results of the system model illustrated in Fig. 8a for different mixed waveform numerologies, using the BLER over increasing frequency spacing Δf as reliability metric. Note that time-synchronized transmission between a narrowband URLLC desired user and two adjacent narrowband interferers is assumed. From this figure, it can be observed that for the same numerology case, CP-OFDM can ensure the optimum performance with a small number of subcarriers being used as a guard band. For the mixed numerology case, the investigated waveforms require larger Δf to achieve equivalent BLER. Although UF-OFDM can provide the best reliability performance in the coexistence of mixed numerology systems but that comes at the expense of an increased computational complexity at the transmitter and receiver side affecting the achievable latency.

We see that several spectral confinement techniques can be applied on top of OFDM taking into account the trade-offs related to spectral efficiency, reliability and latency. However, the exact technique to be used is of an implementation choice and should be transparent to the receiver.

VI. SYSTEM LEVEL SIMULATIONS

The introduced radio components and methods later on interact in a network which makes it interesting to study their performance on system level. For this, we consider a factory hall deployment where the communication is organized in cells and for each cell there is a base station coordinating uplink and downlink traffic.

Scheduling decisions are taken at the base station and are based on the local packet queues and instantaneous channel states which we assume to be known at the base station. A channel state for device j at resource block n is given by the Signal-to-Interference-and-Noise Ratio (SINR) $\gamma_{j,n}(t)$ which is defined as:

$$\gamma_{j,n}(t) = \frac{p_n^s g_{j,n}^s(t)}{\sum_{i=1}^I p_n^i g_{j,n}^i(t) + N_0} \quad (1)$$

Here, p_n^s and p_n^i denote the transmission powers for the signal of interest and the interfering one. The instantaneous

Simulation parameters	Value
Factory hall size	36 m \times 36 m
Deployment grid	3 \times 3 cells
Frequency	5.2 GHz
Frequency reuse factor	1
Packet size	100 bits
Packet arrival rate	10 ms
Scheduling algorithm	Earliest deadline first
OFDM numerology	60 kHz subcarrier spacing
Size of scheduling interval	0.25 ms
Resource block (RB) size	12 subcarriers
Total bandwidth	100 RB
Path loss model	Industrial indoor channel model [24]

Table II: Parameters for system level evaluations

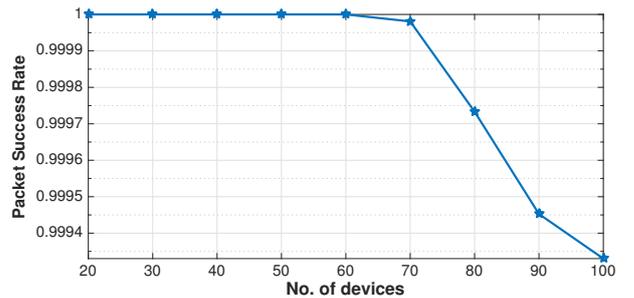


Figure 9: Probability that packets of 100 bits are received in less than 1 ms.

channel gains for the signal of interest and interfering signal are denoted by $g_{j,n}^s$ and $g_{j,n}^i$, respectively. We assume the channel gains to be exponentially distributed with means \bar{g}_j^s and \bar{g}_j^i denoting the corresponding large-scale fading parameters.

Local queues in the scheduler of the base station are organized for both uplink and downlink transmissions, and the packets are queued according to the *earliest deadline first* algorithm. Uplink and downlink transmissions operate in non-overlapping frequency bands and can simultaneously occur at the same time. Given a certain SINR realization $\gamma_{j,n}(t)$, the scheduler assigns to the packet at head of queue as many unscheduled resource blocks as needed to transmit the packet with the required reliability and the resources with the best SINR realization are picked up first. The scheduler repeats until all the available resource blocks at the current scheduling interval are used up.

A modulation and coding scheme m is applied to a resource block if the current SINR realization $\gamma_{j,\min}(t)$ is within the range $\Gamma_m \leq \gamma_{j,\min}(t) < \Gamma_{m+1}$. The SINR thresholds Γ_m are extrapolated for a BLER of 10^{-9} considering eight antennas at the base station and two antennas at the device side with selection combining and uncorrelated fading branches as discussed in Section III. Furthermore, we also use the waveform numerology and frame structure design in line to our proposals in Section IV. The rest of simulation parameters are listed in Table II.

Since a centralized network is considered, the devices in a cell communicate with other devices also inside the

corresponding cell through the base station. From our system evaluations, we estimate the time it takes from the source to arrive to its destination. Based on these data, we compute the probability that a packet is received in less than 1 ms with BLER of 10^{-9} . We call it as *packet success rate* and the corresponding results are given in Figure 9. We can observe that in scenarios with up to 60 devices per cell, no packets are dropped due to deadline expiration. It gives us an idea on the maximum number of devices which can be supported in a certain network deployment and the requirement on the system bandwidth, i.e. 72 MHz in our case to support upto 60 devices per cell.

Thus, by considering the typical industrial scenario as mentioned in [3], we can conclude from our system level evaluations that highly reliable communication can be achieved within the latency bounds by an appropriate radio design.

VII. CONCLUSIONS

Ultra-reliable and low-latency communication can enable many new 5G use cases with challenging requirements such as wireless industrial automation. In this paper, we have comprehensively discussed different radio interface concepts impacting the transmission reliability and the end-to-end communication latency. Diversity is considered to be an important tool to enable reliable wireless communication and can be exploited at different layers of the protocol stack. Performance of low delay tail-biting convolutional codes were analyzed and the basis for link-to-system model was developed to be used for system-level evaluations. Furthermore, suitable numerologies and frame structure design principles for low-latency communication were discussed. It is shown that numerology with 60 kHz subcarrier spacing is an optimum choice for a typical industrial deployments given the tight latency constraints. Coexistence of different applications is the key aspect for 5G networks which require some spectral confinement technique, such as filtering, on top of traditional OFDM. At the end, we performed system level evaluations for a typical industrial deployment using the developed link-to-system model. Based on the simulation results, it has been shown that it is feasible to meet the most strict URLLC requirements with sufficient coverage given a suitable radio design.

ACKNOWLEDGMENT

This work was partially funded by the German Federal Ministry of Education and Research (BMBF) within the project “Koordinierte Industriekommunikation (KoI)”[25]. We appreciate the support from the partners in the project consortium.

REFERENCES

- [1] E. Dahlmann, G. Mildh, and S. Parkvall, “5G Wireless Access: Requirements and Realization,” in *IEEE Communications Magazine*, vol. 52, no. 12, pp. 42–47, Dec. 2014.
- [2] 3GPP TR 38.913 (v14.0.0), “Technical Specification Group Radio Access Network; Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14),” Oct. 2016.
- [3] B. Holfeld, D. Wieruch, T. Wirth, L. Thiele, S. Ashraf, J. Huschke, I. Aktas, and J. Ansari, “Wireless communication for factory automation: An opportunity for LTE and 5G systems,” in *IEEE Communications Magazine*, pp. 36–43, June 2016.
- [4] S. Ashraf, I. Aktas, E. Eriksson, K. Helmersson, and J. Ansari, “Ultra-Reliable and Low-Latency Communication for Wireless Factory Automation: From LTE to 5G,” in *Proc. of IEEE ETFA*, Sept. 2016.
- [5] P. Popovski, “Ultra-reliable communication in 5g wireless systems,” in *1st International Conference on 5G for Ubiquitous Connectivity*, Nov. 2014.
- [6] M. Weiner, M. Jorgovanovic, A. Sahai, and B. Nikolic, “Design of a low-latency, high-reliability wireless communication system for control applications,” in *Proc. of IEEE ICC*, June 2014.
- [7] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, “The 5G-Enabled Tactile Internet: Applications, Requirements, and Architecture,” in *Proc. of IEEE WCNC*, April 2016.
- [8] S. Ashraf, F. Lindqvist, B. Lindoff, and R. Baldemair, “Control channel design trade-offs for ultra-reliable and low-latency communication system,” in *IEEE Globecom Workshop on Ultra-Low Latency and Ultra-High Reliability in Wireless Communication*, Dec. 2015.
- [9] N. Johansson, Y.-P. E. Wang, E. Eriksson, and M. Hessler, “Radio access for ultra-reliable and low-latency 5g communications,” in *IEEE ICC Workshop on 5G and Beyond - Enabling Technologies and Applications*, June 2015.
- [10] B. Holfeld, D. Wieruch, L. Raschkowski, T. Wirth, C. Pallasch, W. Herfs, and C. Brecher, “Radio Channel Characterization at 5.85 GHz for Wireless M2M Communication of Industrial Robots,” in *Proc. of IEEE WCNC*, April 2016.
- [11] D. Wieruch, B. Holfeld, and T. Wirth, “Wireless Factory Automation: Radio Channel Evolution in Repeated Manufacturing Processes,” in *Proc. of ITG Workshop on Smart Antennas (WSA)*, March 2016.
- [12] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior,” in *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [13] V. N. Swamy, S. Suri, P. Rigge, M. Weiner, G. Ranade, A. Sahai, and B. Nikolic, “Cooperative communication for high-reliability low-latency wireless control,” in *Proc. of IEEE ICC*, June 2015.
- [14] V. N. Swamy, P. Rigge, G. Ranade, A. Sahai, and B. Nikolic, “Network coding for high-reliability low-latency wireless control,” in *Proc. of IEEE WCNC*, April 2016.
- [15] M. Serror, C. Dombrowski, K. Wehrle, and J. Gross, “Channel Coding Versus Cooperative ARQ: Reducing Outage Probability in Ultra-Low Latency Wireless Communications,” in *IEEE GLOBECOM Workshops on Ultra-Low Latency and Ultra-High Reliability in Wireless Communication*, Dec. 2015.
- [16] M. Serror, Y. Hu, C. Dombrowski, K. Wehrle, and J. Gross, “Performance Analysis of Cooperative ARQ Systems for Wireless Industrial Networks,” in *Proc. of IEEE WoWMoM*, June 2016.
- [17] E. Malkamaki and H. Leib, “Coded Diversity on Block-fading Channels,” in *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.
- [18] R1 071323 Motorola, “Performance of convolutional codes for the E-UTRA DL control channel,” 3GPP TSG RAN WG1 48bis, 2007.
- [19] R1 073033 Ericsson, “Complexity and performance improvement for convolutional coding,” 3GPP TSG RAN WG1 49bis, 2007.
- [20] Y.-P. Wang and R. Ramesh, “To bite or not to bite - a study of tail bits versus tail-biting,” in *Seventh IEEE PIMRC*, Oct. 1996.
- [21] Y.-P. Wang, R. Ramesh, A. Hassan, and H. Koorapaty, “On MAP Decoding for Tail-Biting Convolutional Codes,” in *Proceedings of IEEE Information Theory Symposium*, 1997.
- [22] R1 163227 Ericsson, “Numerology for NR,” 3GPP TSG RAN WG1 84bis, 2016.
- [23] G. Wunder, P. Jung, M. Kasparick, T. Wild *et al.*, “5GNOW: Non-Orthogonal, Asynchronous Waveforms for Future Mobile Applications,” in *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, 2014.
- [24] E. Tanghe, W. Joseph, L. Verloock, L. Martens, H. Capoen, K. Van Herwegen, and W. Vantomme, “The Industrial Indoor Channel: Large-Scale and Temporal Fading at 900, 2400, and 5200 MHz,” in *IEEE Transactions on Wireless Communication*, vol. 7, no. 7, pp. 2740–2751, July 2008.
- [25] “Koordinierte Industriekommunikation (KoI),” <http://koi-projekt.de/>.